Optimal quantum detectors for unambiguous detection of mixed states

Yonina C. Eldar*
Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel

Mihailo Stojnic† and Babak Hassibi‡
Department of Electrical Engineering, California Institute of Technology, Pasadena, California 91125, USA

(Received 7 December 2003; published 21 June 2004)

We consider the problem of designing an optimal quantum detector that distinguishes unambiguously between a collection of mixed quantum states. Using arguments of duality in vector space optimization, we derive necessary and sufficient conditions for an optimal measurement that minimizes the probability of an inconclusive result. We show that the previous optimal measurements that were derived for certain special cases satisfy these optimality conditions. We then consider state sets with strong symmetry properties, and show that the optimal measurement operators for distinguishing between these states share the same symmetries, and can be computed very efficiently by solving a reduced size semidefinite program.

DOI: 10.1103/PhysRevA.69.062318 PACS number(s): 03.67.Hk

I. INTRODUCTION

The problem of detecting information stored in the state of a quantum system is a fundamental problem in quantum information theory. Several approaches have emerged to distinguishing between a collection of nonorthogonal quantum states. In one approach, a measurement is designed to minimize the probability of a detection error [1–10]. A more recent approach, referred to as unambiguous detection [11–19], is to design a measurement that with a certain probability returns an inconclusive result, but such that if the measurement returns an answer, then the answer is correct with probability 1. An interesting alternative approach for distinguishing between a collection of quantum states, which is a combination of the previous two approaches, is to allow for a certain probability of an inconclusive result, and then maximize the probability of correct detection [19–21].

We consider a quantum state ensemble consisting of m density operators $\{\rho_i, 1 \leqslant i \leqslant m\}$ on an n-dimensional complex Hilbert space $\mathcal{H}$, with prior probabilities $\{p_i > 0, 1 \leqslant i \leqslant m\}$. A pure-state ensemble is one in which each density operator ρ_i is a rank-one projector $|\phi_i\rangle\langle\phi_i|$, where the vectors $|\phi_i\rangle$, though evidently normalized to unit length, are not necessarily orthogonal. Our problem is to design a quantum detector to distinguish unambiguously between the states $\{\rho_i\}$.

Chefles [16] showed that a necessary and sufficient condition for the existence of unambiguous measurements for distinguishing between a collection of *pure* quantum states is that the states are linearly independent. Necessary and sufficient conditions on the optimal measurement minimizing the probability of an inconclusive result for pure states were derived in Ref. [18]. The optimal measurement when distinguishing between a broad class of symmetric pure-state sets was also considered in Ref. [18].

The problem of unambiguous detection between *mixed* state ensembles has received considerably less attention. Rudolph, Speakkens, and Turner [22] showed that unambiguous detection between mixed quantum states is possible as long as one of the density operators in the ensemble has a nonzero overlap with the intersection of the kernels of the other density operators. They then consider the problem of unambiguous detection between two mixed quantum states, and derive upper and lower bounds on the probability of an inconclusive result. They also develop a closed form solution for the optimal measurement in the case in which both states have kernels of dimension 1. In Ref. [23], the authors consider the problem of unambiguous discrimination between two general density matrices.

In this paper we develop a general framework for unambiguous state discrimination between a collection of mixed quantum states, which can be applied to any number of states with arbitrary prior probabilities. For our measurement we consider general positive operator-valued measures [2,24], consisting of $m+1$ measurement operators. We derive a set of necessary and sufficient conditions for an optimal measurement that minimizes the probability of an inconclusive result, by exploiting principles of duality theory in vector space optimization. We then show that the previous optimal measurements that were derived for certain special cases satisfy these optimality conditions.

Next, we consider geometrically uniform (GU) and compound GU (CGU) state sets [7,8,25], which are state sets with strong symmetry properties. We show that the optimal measurement operators for unambiguous discrimination between such state sets are also GU and CGU, respectively, with generators that can be computed very efficiently by solving a reduced size semidefinite program.

The paper is organized as follows. In Sec. II, we provide a statement of our problem. In Sec. III we develop the necessary and sufficient conditions for optimality using Lagrange duality theory. Some special cases are considered in Sec. IV. In Sec. V we consider the problem of distinguishing between a collection of states with a broad class of symmetry properties.

*Electronic address: yonina@ee.technion.ac.il
†Electronic address: mihailo@systems.caltech.edu
‡Electronic address: hassibi@systems.caltech.edu

II. PROBLEM FORMULATION

Assume that a quantum channel is prepared in a quantum state drawn from a collection of mixed states, represented by density operators $\{\rho_i, 1 \leq i \leq m\}$ on an n-dimensional complex Hilbert space $\mathcal{H}$. We assume without loss of generality that the eigenvectors of $\rho_i, 1 \leq i \leq m$, collectively span [33] $\mathcal{H}$.

To detect the state of the system a measurement is constructed comprising $m+1$ measurement operators $\{\Pi_i, 0 \leq i \leq m\}$ that satisfy

$$\sum_{i=0}^{m} \Pi_i = I, \tag{1}$$

$\Pi_i \geq 0, 0 \leq i \leq m$. The measurement operators are constructed so that either the state is correctly detected, or the measurement returns an inconclusive result. Thus, each of the operators $\Pi_i, 1 \leq i \leq m$ correspond to detection of the corresponding states $\rho_i, 1 \leq i \leq m$, and Π_0 corresponds to an inconclusive result.

Given that the state of the system is ρ_j, the probability of obtaining outcome i is $\mathrm{Tr}(\rho_j \Pi_i)$. Therefore, to ensure that each state is either correctly detected or an inconclusive result is obtained, we must have

$$\mathrm{Tr}(\rho_j \Pi_i) = \eta_i \delta_{ij}, \ 1 \leq i,j \leq m, \tag{2}$$

for some $0 \leq \eta_i \leq 1$. Since from Eq. (1), $\Pi_0 = I - \Sigma_{i=1}^{m} \Pi_i$, Eq. (2) implies that $\mathrm{Tr}(\rho_i \Pi_0) = 1 - \eta_i$, so that given that the state of the system is ρ_i, the state is correctly detected with probability η_i, and an inconclusive result is returned with probability $1 - \eta_i$.

It was shown in Ref. [16] that for pure-state ensembles consisting of rank-one density operators $\rho_i = |\phi_i\rangle\langle\phi_i|$, Eq. (2) can be satisfied if and only if the vectors $|\phi_i\rangle$ are linearly independent. For mixed states, it was shown in Ref. [22] that Eq. (2) can be satisfied if and only if one of the density operators ρ_i has a nonzero overlap with the intersection of the kernels of the other density operators. Specifically, denote by $\mathcal{K}_i$ the null space of ρ_i and let

$$\mathcal{S}_i = \cap_{j=1, j\neq i}^{m} \mathcal{K}_j \tag{3}$$

denote the intersection of $\mathcal{K}_j, 1 \leq j \leq m, j \neq i$. Then to satisfy Eq. (2) the eigenvectors of Π_i must be contained in $\mathcal{S}_i$ and must not be entirely contained in $\mathcal{K}_i$. This implies that $\mathcal{K}_i$ must not be entirely contained in $\mathcal{S}_i$. Some examples of mixed states for which unambiguous detection is possible are given in Ref. [22].

If the state ρ_i is prepared with prior probability p_i, then the total probability of correctly detecting the state is

$$P_D = \sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i \Pi_i). \tag{4}$$

Our problem therefore is to choose the measurement operators Π_i, $0 \leq i \leq m$ to maximize P_D, subject to the constraints (1) and

$$\mathrm{Tr}(\rho_j \Pi_i) = 0, \quad 1 \leq i,j \leq m, \quad i \neq j. \tag{5}$$

To satisfy Eq. (5), Π_i must lie in $\mathcal{S}_i$ defined by Eq. (3), so that

$$\Pi_i = P_i \Pi_i P_i, \quad 1 \leq i \leq m, \tag{6}$$

where P_i is the orthogonal projection onto $\mathcal{S}_i$. Denoting by Θ_i an $n \times r$ matrix whose columns form an arbitrary orthonormal basis for $\mathcal{S}_i$ where $r = \dim(\mathcal{S}_i)$, we can express P_i as $P_i = \Theta_i \Theta_i^*$. From Eqs. (6) and (1) we then have that

$$\Pi_i = \Theta_i \Delta_i \Theta_i^*, \quad 1 \leq i \leq m, \tag{7}$$

where $\Delta_i = \Theta_i^* \Pi_i \Theta_i$ is an $r \times r$ matrix satisfying

$$\sum_{i=1}^{m} \Theta_i \Delta_i \Theta_i^* \leq I, \tag{8}$$

$\Delta_i \geq 0, 1 \leq i \leq m$. Therefore, our problem reduces to maximizing

$$P_D = \sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i \Theta_i \Delta_i \Theta_i^*), \tag{9}$$

subject to Eq. (8).

To show that the problem of Eqs. (9) and (8) does not depend on the choice of orthonormal basis Θ_i, we note that any orthonormal basis for $\mathcal{S}_i$ can be expressed as the columns of Ψ_i, where $\Psi_i = \Theta_i U_i$ for some $r \times r$ unitary matrix U_i. Substituting Ψ_i instead of Θ_i in Eqs. (9) and (8), our problem becomes that of maximizing

$$P_D = \sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i \Psi_i \Delta_i \Psi_i^*) = \sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i \Theta_i \Delta_i' \Theta_i^*), \tag{10}$$

where $\Delta_i' = \mathrm{U}_i \Delta_i U_i^*$, subject to

$$\sum_{i=1}^{m} \Psi_i \Delta_i \Psi_i^* = \sum_{i=1}^{m} \Theta_i \Delta_i' \Theta_i^* \leq I, \tag{11}$$

$\Delta_i \geq 0, 1 \leq i \leq m$. Since $\Delta_i \geq 0$ if and only if $\Delta_i' \geq 0$, the problem of Eqs. (10) and (11) is equivalent to that of Eqs. (9) and (8).

Equipped with the standard operations of addition and multiplication by real numbers, the space $\mathcal{B}$ of all Hermitian $n \times n$ matrices is an n^2-dimensional *real* vector space. As noted in Ref. [22], by choosing an appropriate basis for $\mathcal{B}$, the problem of maximizing P_D subject to Eq. (8) can be put in the form of a standard semidefinite programming problem, which is a convex optimization problem; for a detailed treatment of semidefinite programming problems see, e.g., Refs. [26–29]. By exploiting the many well-known algorithms for solving semidefinite programs [29], e.g., interior point methods [26,28,34], the optimal measurement can be computed very efficiently in polynomial time within any desired accuracy.

Using elements of duality theory in vector space optimization, in the following section we derive necessary and sufficient conditions on the measurement operators $\Pi_i = \Theta_i \Delta_i \Theta_i^*$ to maximize P_D of Eq. (9) subject to Eq. (8).

III. CONDITIONS FOR OPTIMALITY

A. Dual-problem formulation

To derive necessary and sufficient conditions for optimality on the matrices Δ_i we first derive a dual problem, using Lagrange duality theory [30].

Denote by Λ the set of all ordered sets $\Pi=\{\Pi_i=\Theta_i\Delta_\iota\Theta_i^*\}_{i=1}^m$ satisfying Eq. (8) and define $J(\Pi)=\Sigma_{i=1}^m p_i\mathrm{Tr}(\rho_i\Theta_i\Delta_i\Theta_i^*)$. Then our problem is

$$\max_{\Pi\in\Lambda} J(\Pi). \tag{12}$$

We refer to this problem as the primal problem, and to any $\Pi\in\Lambda$ as a primal feasible point. The optimal value of $J(\Pi)$ is denoted by $\hat{J}$.

To develop the dual problem associated with Eq. (12) we first compute the Lagrange dual function, which is given by

$$g(Z)$$

$$= \min_{\Delta_i\geqslant 0}\left\{-\sum_{i=1}^m p_i\mathrm{Tr}(\rho_i\Theta_i\Delta_i\Theta_1^*) + \mathrm{Tr}\left[Z\left(\sum_{i=1}^m \Theta_i\Delta_i\Theta_i^* - I\right)\right]\right\}$$

$$=\min_{\Delta_i\geqslant 0}\left\{\sum_{i=1}^m \mathrm{Tr}[\Delta_i\Theta_i^*(Z-p_i\rho_i)\Theta_i] - \mathrm{Tr}(Z)\right\}, \tag{13}$$

where $Z\geqslant 0$ is the Lagrange dual variable. Since $\Delta_i\geqslant 0$, $1\leqslant i\leqslant m$, we have that $\mathrm{Tr}(\Delta_i X)\geqslant 0$ for any $X\geqslant 0$. If X is not positive semidefinite, then we can always choose Δ_i such that $\mathrm{Tr}(\Delta_i X)$ is unbounded below. Therefore,

$$g(Z)=\begin{cases}-\mathrm{Tr}(Z), & A_i\geqslant 0, 1\leqslant i\leqslant m, Z\geqslant 0\\ -\infty, & \text{otherwise},\end{cases} \tag{14}$$

where

$$A_i=\Theta_i^*(Z-p_i\rho_i)\Theta_i, \quad 1\leqslant i\leqslant m. \tag{15}$$

It follows that the dual problem associated with Eq. (12) is

$$\min_{z}\mathrm{Tr}(Z) \tag{16}$$

subject to

$$\Theta_i^*(Z-p_i\rho_i)\Theta_i\geqslant 0, \quad 1\leqslant i\leqslant m,$$

$$Z\geqslant 0. \tag{17}$$

Denoting by Γ the set of all Hermitian operators Z such that $\Theta_1^*(Z-p_i\rho_i)\Theta_i\geqslant 0, 1\leqslant i\leqslant m$, and $Z\geqslant 0$, and defining $T(Z)=\mathrm{Tr}(Z)$, the dual problem can be written as

$$\min_{Z\in\Gamma} T(Z). \tag{18}$$

We refer to any $Z\in\Gamma$ as a dual feasible point. The optimal value of $T(Z)$ is denoted by $\hat{T}$.

B. Optimality conditions

We can immediately verify that both the primal and the dual problem are strictly feasible. Therefore, their optimal values are attainable and the duality gap is zero [29], i.e.,

$$\hat{J}=\hat{T}. \tag{19}$$

In addition, for any $\Pi=\{\Pi_i=\Theta_i\Delta_i\Theta_i^*\}_{i=1}^m\in\Lambda$, and $Z\in\Gamma$,

$$T(Z)-J(\Pi)=\mathrm{Tr}\left(\sum_{i=1}^m\Theta_i\Delta_i\Theta_i^*(Z-p_i\rho_i)+\Pi_0 Z\right)\geqslant 0, \tag{20}$$

where $\Pi_0=I-\Sigma_{i=1}^m\Theta_i\Delta_i\Theta_1^*\geqslant 0$. Note that Eq. (20) can be used to develop an upper bound on the optimal probability of correct detection $\hat{J}$. Indeed, since for any $Z\in\Gamma, T(Z)\geqslant J(\Pi)$, we have that $\hat{J}\leqslant T(Z)$ for any dual feasible Z.

Now, let $\hat{\Pi}_i=\Theta_i\hat{\Delta}_i\Theta_i^*$, $1\leqslant i\leqslant m$ and $\hat{\Pi}_0=I-\Sigma_{i=1}^m\hat{\Pi}_i$ denote the optimal measurement operators that maximize Eq. (9) subject to Eq. (8), and let $\hat{Z}$ denote the optimal Z that minimizes Eq. (16) subject to Eq. (17). From Eqs. (19) and (20) we conclude that

$$\mathrm{Tr}\left(\sum_{i=1}^m\hat{\Pi}_i\Theta_i^*(\hat{Z}-p_i\rho_i)\Theta_i+\hat{\Pi}_0\hat{Z}\right)=0. \tag{21}$$

Since $\hat{\Delta}_i\geqslant 0, \hat{Z}\geqslant 0$, and $\Theta_i^*(\hat{Z}-p_i\rho_i)\Theta_i\geqslant 0$, $1\leqslant i\leqslant m$, Eq. (21) is satisfied if and only if

$$\hat{Z}\hat{\Pi}_0=0, \tag{22}$$

$$\Theta_i^*(\hat{Z}-p_i\rho_i)\Theta_i\hat{\Delta}_i=0, \quad 1\leqslant i\leqslant m. \tag{23}$$

Once we find the optimal $\hat{Z}$ that minimizes the dual problem (16), the constraints (22) and (23) are necessary and sufficient conditions on the optimal measurement operators $\hat{\Pi}_i$. We have already seen that these conditions are necessary. To show that they are sufficient, we note that if a set of feasible measurement operators $\hat{\Pi}_i$ satisfies Eqs. (22) and (23), then $\mathrm{Tr}[\Sigma_{i=1}^m\hat{\Delta}_i\Theta_i^*(\hat{Z}-p_i\rho_i)\Theta_i+\hat{\Pi}_0\hat{Z}]=0$ so that from Eq. (20), $J(\hat{\Pi})=T(\hat{Z})=\hat{J}$.

We summarize our results in the following theorem.

Theorem 1. Let $\{\rho_{i,}1\leqslant i\leqslant m\}$ denote a set of density operators with prior probabilities $\{p_i>0, 1\leqslant i\leqslant m\}$, and let $\{\Theta_i, 1\leqslant i\leqslant m\}$ denote a set of matrices such that the columns of Θ_i form on orthonormal basis for $\mathcal{S}_i=\cap_{j=1,j\neq i}^m\mathcal{K}_j$, where $\mathcal{K}_i$ the null space of ρ_i. Let Λ denote the set of all ordered sets of Hermitian measurement operators $\Pi=\{\Pi_i\}_{i=0}^m$ that satisfy $\Pi_i\geqslant 0, \Sigma_{i=0}^m\Pi_i=I$, and $\mathrm{Tr}(\rho_j\Pi_i)=0$, $1\leqslant i\leqslant m$, $i\neq j$ and let Γ denote the set of Hermitian matrices Z such that $Z\geqslant 0$, $\Theta_1^*(Z-p_i\rho_i)\Theta_i$, $1\leqslant i\leqslant m$. Consider the problem max $_{\Pi\in\Lambda}J(\Pi)$ and the dual problem min $_{Z\in\Gamma}T(Z)$, where $J(\Pi)=\Sigma_{i=1}^m p_i\mathrm{Tr}(\rho_i\Pi_i)$ and $T(Z)=\mathrm{Tr}(Z)$. Then we have the following.

(1) For any $Z\in\Gamma$ and $\Pi\in\Lambda$, $T(Z)\geqslant J(\Pi)$.

(2) There is an optimal Π, denoted $\hat{\Pi}$, such that $\hat{J} = J(\hat{\Pi}) \geqslant J(\Pi)$ for any $\Pi \in \Lambda$.

(3) There is an optimal Z, denoted $\hat{Z}$ and such that $\hat{T} = T(\hat{Z}) \leqslant T(Z)$ for any $Z \in \Gamma$.

(4) $\hat{T} = \hat{J}$.

(5) Necessary and sufficient conditions on the optimal measurement operators $\hat{\Pi}_i$ are that there exists a $Z \in \Gamma$ such that

$$Z\hat{\Pi}_0 = 0, \qquad (24)$$

$$\Theta_i^*(Z - p_i\rho)\Theta_i\hat{\Delta}_i = 0, \quad 1 \leqslant i \leqslant m, \qquad (25)$$

where $\hat{\Pi}_i = \Theta_i\hat{\Delta}\Theta_i^*, 1 \leqslant i \leqslant m$, and $\hat{\Delta}_i \geqslant 0$.

(6) Given $\hat{Z}$, necessary and sufficient conditions on the optimal measurement operators $\hat{\Pi}_i$ are

$$\hat{Z}\hat{\Pi}_0 = 0, \qquad (26)$$

$$\Theta_i^*(\hat{Z} - p_i\rho_i)\Theta_i\hat{\Delta}_i = 0, \quad 1 \leqslant i \leqslant m. \qquad (27)$$

Although the necessary and sufficient conditions of Theorem 1 are hard to solve, they can be used to verify a solution and to gain some insight into the optimal measurement operators. In the following section we show that the previous optimal measurements that were derived in the literature for certain special cases satisfy these optimality conditions.

IV. SPECIAL CASES

We now consider two special cases that were addressed in Ref. [22], for which a closed form solution exists. In Sec. IV A we consider the case in which the spaces $\mathcal{S}_i$ defined by Eq. (3) are orthogonal, and in Sec. IV B we consider the problem of distinguishing unambiguously between two density operators with $\dim(\mathcal{S}_i) = 1, 1 \leqslant i \leqslant 2$.

A. Orthogonal null spaces $\mathcal{S}_i$

The first case we consider is the case in which the null spaces $\mathcal{S}_i$ are orthogonal, so that

$$P_iP_j = \delta_{ij}, \quad 1 \leqslant i,j \leqslant m, \qquad (28)$$

where P_i is an orthogonal projection onto $\mathcal{S}_i$. It was shown in Ref. [22] that in this case the optimal measurement operators are

$$\hat{\Pi}_i = P_i = \Theta_i\Theta_i^*, \quad 1 \leqslant i \leqslant m. \qquad (29)$$

In Appendix A we show that the optimal solution of the dual problem can be expressed as

$$\hat{Z} = \sum_{i=1}^{m} p_iP_i\rho_iP_i. \qquad (30)$$

It can easily be shown that $\hat{Z}$ and $\hat{\Pi}_i$ of Eqs. (30) and (29) satisfy the optimality conditions of Theorem 1.

B. Null spaces of dimension 1

We now consider the case of distinguishing between two density operators ρ_1 and ρ_2, where $\mathcal{S}_1$ and $\mathcal{S}_2$ both have dimension equal to 1. In this case, as we show in Appendix B, the optimal dual solution is

$$\hat{Z} = \begin{cases} d_1P_1, & d_2 - d_1|f|^2 \leqslant 0 \\ d_2P_2, & d_1 - d_2|f|^2 \leqslant 0 \\ d_2(\Theta_2 + s\Theta_2^{\perp})(\Theta_2 + s\Theta_2^{\perp})^*, & \text{otherwise}, \end{cases} \qquad (31)$$

where P_i is an orthogonal projection onto $\mathcal{S}_i$, $\Theta_2^{\perp}$ is a unit norm vector in the span of Θ_1 and Θ_2 such that $\Theta_2^*\Theta_2^{\perp} = 0$, and

$$d_i = p_i\Theta_i^*p_i\Theta_i, \quad 1 \leqslant i \leqslant 2,$$

$$s = \frac{f^*}{e^*}\left(\sqrt{\frac{d_i}{d_2|f|^2}} - 1\right),$$

$$f = \Theta_2^*\,\Theta_1,$$

$$e = (\Theta_2^{\perp})^*\Theta_1. \qquad (32)$$

The optimal measurement operators for this case were developed in Ref. [22], and can be written as

$$\{\hat{\Pi}_i\}_{i=1}^2 = \begin{cases} \hat{\Pi}_1 = P_1, \hat{\Pi}_2 = 0, & d_2 - d_1|f|^2 \leqslant 0 \\ \hat{\Pi}_1 = 0, \hat{\Pi}_2 = P_2, & d_1 - d_2|f|^2 \leqslant 0 \\ \hat{\Pi}_1 = \alpha_1P_1, \hat{\Pi}_2 = \alpha_2P_2, & \text{otherwise}, \end{cases} \qquad (33)$$

where

$$\alpha_1 = \frac{1 - \sqrt{\frac{d_2|f|^2}{d_1}}}{1 - |f|^2},$$

$$\alpha_2 = \frac{1 - \sqrt{\frac{d_1|f|^2}{d_2}}}{1 - |f|^2}. \qquad (34)$$

We now show that $\hat{Z}$ and $\hat{\Pi}$ of Eqs. (31) and (33) satisfy the optimality conditions of Theorem 1. To this end we note that from Eq. (33),

$$\{\hat{\Delta}_i\}_{i=1}^2 = \begin{cases} \hat{\Delta}_1 = 1, \hat{\Delta}_2 = 0, & d_2 - d_1|f|^2 \leqslant 0 \\ \hat{\Delta}_1 = 0, \hat{\Delta}_2 = 1, & d_1 - d_2|f|^2 \leqslant 0 \\ \hat{\Delta}_1 = \alpha_1, \hat{\Delta}_2 = \alpha_2, & \text{otherwise}. \end{cases} \qquad (35)$$

From Eqs. (31)–(35) we have that if $d_2 - d_1|f|^2 \leqslant 0$, then

$$\Theta_1^*(\hat{Z} - p_1\rho_1)\Theta_1\hat{\Delta}_1 = d_1 - \Theta_1^*p_1\rho_1\Theta_1 = 0,$$

$$\Theta_2^*(\hat{Z} - p_2\rho_2)\Theta_2\hat{\Delta}_2 = 0,$$

$$\hat{Z}\hat{\Pi}_0 = \hat{Z}(I - \hat{\Pi}_1) = d_1\Theta_1\Theta_1^* - d_1\Theta_1\Theta_1^* = 0. \tag{36}$$

Similarly, if $d_1 - d_2|f|^2 \leqslant 0$, then

$$\Theta_1^*(\hat{Z} - p_1\rho_1)\Theta_1\hat{\Delta}_1 = 0,$$

$$\Theta_2^*(\hat{Z} - p_2\rho_2)\Theta_2\hat{\Delta}_2 = d_2 - \Theta_2^* p_2\rho_2\Theta_2 = 0,$$

$$\hat{Z}\hat{\Pi}_0 = \hat{Z}(I - \hat{\Pi}_2) = d_2\Theta_2\Theta_2^* - d_2\Theta_2\Theta_2^* = 0. \tag{37}$$

Finally, if neither of the conditions $d_1 - d_2|f|^2 \leqslant 0$, $d_2 - d_1|f|^2 \leqslant 0$ hold, then

$$\begin{aligned}\Theta_1^*(\hat{Z} - p_1\rho_1\Theta_1\hat{\Delta}_1) &= [d_2(f^* + e^*s)(f^* + e^*s)^* - d_1]\frac{1 - \sqrt{\dfrac{d_2|f|^2}{d_1}}}{1 - |f|^2} \\ &= \left[d_2|f|^2\left(\sqrt{\frac{d_1}{d_2|f|^2}}\right)^2 - d_1\right]\frac{1 - \sqrt{\dfrac{d_2|f|^2}{d_1}}}{1 - |f|^2} \\ &= 0,\end{aligned} \tag{38}$$

$$\Theta_2^*(\hat{Z} - p_2\rho_2)\Theta_2\hat{\Delta}_2 = (\Theta_2^*\hat{Z}\Theta_2 - d_2)\frac{1\sqrt{\dfrac{d_1|f|^2}{d_2}}}{1 - |f|^2} = 0, \tag{39}$$

and

$$\hat{Z}\hat{\Pi}_0 = \hat{Z} - \hat{Z}\hat{\Pi}_1 - \hat{Z}\hat{\Pi}_2 = \hat{Z} - \hat{\Delta}_1\hat{Z}\Theta_1\Theta_1^* - \hat{\Delta}_2\hat{Z}\Theta_2\Theta_2^*. \tag{40}$$

To show that $\hat{Z}\hat{\Pi}_0 = 0$, we note that

$$\begin{aligned}\hat{Z}\Theta_1\Theta_1^* = {}& d_2(|f|^2 + s^*ef^*)\Theta_2\Theta_2^* + d_2(s|f|^2 + ss^*ef^*)\Theta_2^{\perp}\Theta_2^* \\ & + d_2(e^*f + s^*|e|^2)\Theta_2\Theta_2^{\perp *} + d_2(se^*f + ss^*|e|^2)\Theta_2^{\perp}\Theta_2^{\perp *},\end{aligned} \tag{41}$$

and

$$\hat{Z}\Theta_2\Theta_2^* = d_2\Theta_2\Theta_2^* + d_2 s\Theta_2^{\perp}\Theta_2^*. \tag{42}$$

Substituting Eqs. (41) and (42) into Eq. (40), and after some algebraic manipulations, we have that

$$\hat{Z}\hat{\Pi}_0 = \hat{Z} - \hat{\Delta}_1\hat{Z}\Theta_1\Theta_1^* - \hat{\Delta}_2\hat{Z}\Theta_2\Theta_2^* = 0. \tag{43}$$

Combining Eqs. (36)–(43) we conclude that the optimal measurement operators of Eq. (22) satisfy the optimality conditions of Theorem 1.

V. OPTIMAL DETECTION OF SYMMETRIC STATES

We now consider the case in which the quantum state ensemble has symmetry properties referred to as GU and GCU. These symmetry properties are quite general, and include many cases of practical interest.

Under a variety of different optimality criteria the optimal measurement for distinguishing between GU and CGU state sets was shown to be GU and CGU, respectively [7,8,18,19]. In particular, it was shown in Ref. [18] that the optimal measurement for unambiguous detection between linearly independent GU and CGU pure states is GU and CGU, respectively. We now generalize this result to unambiguous detection of mixed GU and CGU state sets.

VI. GU STATE SETS

A GU state set is defined as a set of density operators $\{\rho_i, 1 \leqslant i \leqslant m\}$ such that $\rho_i = U_i\rho U_1^*$ where ρ is an arbitrary *generating operator* and the matrices $\{U_i, 1 \leqslant i \leqslant m\}$ are unitary and form an Abelian group $\mathcal{G}$ [8,31]. For concreteness, we assume that $U_1 = I$.

The group $\mathcal{G}$ is the *generating group* of $\mathcal{S}$. For consistency with the symmetry of $\mathcal{S}$, we will assume equiprobable prior probabilities on $\mathcal{S}$.

As we now show, the optimal measurement operators that maximize the probability of correct detection when distinguishing unambiguously between the density operators of a GU state set are also GU with the same generating group. The corresponding generator can be computed very efficiently in polynomial time.

Suppose that the optimal measurement operators that maximize

$$J(\{\Pi_i\}) = \sum_{i=1}^{m} \mathrm{Tr}(\rho_i\Pi_i) \tag{44}$$

subject to Eqs. (8) and (5) are $\hat{\Pi}_i$, and let $\hat{J} = J(\{\hat{\Pi}_i\}) = \sum_{i=1}^m \mathrm{Tr}(\rho_i\hat{\Pi}_i)$. Let $r(j,i)$ be the mapping from $\mathcal{I}\times\mathcal{I}$ to $\mathcal{I}$ with $\mathcal{I} = \{1, \ldots, m\}$, defined by $r(j,i) = k$ if $U_j^*U_i = U_k$. Then the measurement operators $\hat{\Pi}_i^{(j)} = U_j\hat{\Pi}_{r(j,i)}U_j^*$ and $\hat{\Pi}_0^{(j)} = I - \sum_{i=1}^m \hat{\Pi}_i^{(j)}$ for any $1 \leqslant j \leqslant m$ are also optimal. Indeed, since $\hat{\Pi}_i \geqslant 0, 1 \leqslant i \leqslant m$ and $\sum_{i=1}^m \hat{\Pi}_i \leqslant I, \hat{\Pi}_i^{(j)} \geqslant 0$, $1 \leqslant i \leqslant m$ and

$$\sum_{i=1}^{m} \hat{\Pi}_i^{(j)} = U_j\left(\sum_{i=1}^{m} \hat{\Pi}_i\right)U_j^* \leqslant U_jU_j^* = I. \tag{45}$$

Using the fact that $\rho_i = U_i\rho U_i^*$ for some generator ρ,

$$\begin{aligned}J(\{\hat{\Pi}_i^{(j)}\}) &= \sum_{i=1}^{m} \mathrm{Tr}(\rho U_j^*U_j\hat{\Pi}_{r(j,i)}U_j^*U_i) \\ &= \sum_{k=1}^{m} \mathrm{Tr}(\rho U_k^*\hat{\Pi}_kU_k) \\ &= \sum_{i=1}^{m} \mathrm{Tr}(\rho_i\hat{\Pi}_i) \\ &= \hat{J}.\end{aligned} \tag{46}$$

Finally, for $l \neq i$,

$$\begin{aligned}\mathrm{Tr}(\rho_l\hat{\Pi}_i^{(j)}) &= \mathrm{Tr}(U_l\rho U_l^* U_j\hat{\Pi}_{r(j,i)}U_j^*)\\ &=\mathrm{Tr}(U_s\rho U_s^*\hat{\Pi}_{r(j,i)})\\ &=\mathrm{Tr}(\rho_s\hat{\Pi}_k)\\ &=0, \end{aligned}\tag{47}$$

where $U_s = U_j^* U_l$ and $U_k = U_j^* U_i$ and the last equality follows from the fact that since $l \neq i$, $s \neq k$.

It was shown in Refs. [8,19] that if the measurement operators $\hat{\Pi}_i^{(j)}$ are optimal for any j, then $\{\overline{\Pi}_i = (1/m)\Sigma_{j=1}^m \hat{\Pi}_i^{(j)}, 1 \leq i \leq m\}$ and $\overline{\Pi}_0 = I - \Sigma_{i=1}^m \overline{\Pi}_i$ are also optimal. Furthermore, $\overline{\Pi}_i = U_i\hat{\Pi}U_i^*$ where $\hat{\Pi} = (1/m)\Sigma_{k=1}^m U_k^*\hat{\Pi}_k U_k$.

We therefore conclude that the optimal measurement operators can always be chosen to be GU with the same generating group $\mathcal{G}$ as the original state set. Thus, to find the optimal measurement operators all we need is to find the optimal generator $\hat{\Pi}$. The remaining operators are obtained by applying the group $\mathcal{G}$ to $\hat{\Pi}$.

Since the optimal measurement operators satisfy $\Pi_i = U_i\Pi U_i^*, 1 \leq i \leq m$ and $\rho_i = U_i\rho U_i^*$, $\mathrm{Tr}(\rho_i\Pi_i) = \mathrm{Tr}(\rho\Pi)$, so that the problem (9) reduces to the maximization problem

$$\max_{\Pi\in\mathcal{B}} \mathrm{Tr}(\rho\Pi),\tag{48}$$

where $\mathcal{B}$ is the set of $n \times n$ Hermitian operators, subject to the constraints

$$\begin{gathered}\Pi \geq 0,\\ \sum_{i=1}^m U_i\Pi U_i^* \leq I,\\ \mathrm{Tr}(\Pi U_i\rho U_i^*) = 0, \quad 2 \leq i \leq m.\end{gathered}\tag{49}$$

The problem of Eqs. (48) and (49) is a (convex) semidefinite programming problem, and therefore the optimal Π can be computed very efficiently in polynomial time within any desired accuracy [26,28,29], for example, using the LMI toolbox on Matlab. Note that the problem of Eqs. (48) and (49) has n^2 real unknowns and $m+1$ constraints, in contrast with the original maximization problem (9) subject to Eqs. (8) and (5) which has mn^2 real unknowns and m^2+1 constraints.

VII. CGU STATE SETS

A CGU state set is defined as a set of density operators $\{\rho_{ik,}\ 1 \leq i \leq l\ 1 \leq k \leq r\}$ such that $\rho_{ik} = U_i\phi_k U_i^*$ for some generating density operators $\{\rho_{k,}\ 1, \leq k \leq r\}$, where the matrices $\{U_i, 1 \leq i \leq l\}$ are unitary and form an Abelian group $\mathcal{G}$ [8,25]. A CGU state set is, in general, not GU. However, for every k, the operators $\{\rho_{ik}, 1 \leq i \leq l\}$ are GU with generating group $\mathcal{G}$.

Using arguments similar to those of Sec. VI and Ref. [19] we can show that the optimal measurement operators corresponding to a CGU state set can always be chosen to be GU with the same generating group $\mathcal{G}$ as the original state set. Thus, to find the optimal measurement operators all we need is to find the optimal generators $\hat{\Pi}_k$. The remaining operators are obtained by applying the group $\mathcal{G}$ to each of the generators $\hat{\Pi}_k$.

Since the optimal measurement operators satisfy $\Pi_{ik} = U_i\Pi_k U_i^*, 1 \leq i \leq l, 1 \leq k \leq r$ and $\rho_{ik} = U_i\rho_k U_i^*, \mathrm{Tr}(\rho_{\mathrm{ik}}\Pi_{ik}) = \mathrm{Tr}(\rho_k\Pi_k)$, so that the problem (9) reduces to the maximization problem

$$\max_{\Pi_k\in B}\sum_{k=1}^r \mathrm{Tr}(\rho_k\Pi_k),\tag{50}$$

subject to the constraints

$$\begin{gathered}\sum_{i=1}^l\sum_{k=1}^r U_i\Pi_k U_i^* \leq I, \quad \Pi_k \geq 0, 1 \leq k \leq r,\\ \mathrm{Tr}(\Pi_k U_i\rho_j U_i^*) = 0, \quad 1 \leq k,j \leq r, 1 \leq i \leq l,\\ \text{if } i = 1 \quad \text{then} \quad k \neq j.\end{gathered}\tag{51}$$

Since this problem is a (convex) semidefinite programming problem, the optimal generators Π_k can be computed very efficiently in polynomial time within any desired accuracy [26,28,29]. Note that the problem of Eqs. (50) and (51) has rn^2 real unknowns and $lr+1$ constraints, in contrast with the original maximization which has lrn^2 real unknowns and $(lr)^2+1$ constraints.

VIII. CONCLUSION

We considered the problem of distinguishing unambiguously between a collection of *mixed* quantum states. Using elements of duality theory in vector space optimization, we derived a set of necessary and sufficient conditions on the optimal measurement operators. We then considered some special cases for which closed form solutions are known, and showed that they satisfy our optimality conditions. We also showed that in the case in which the states to be distinguished have strong symmetry properties, the optimal measurement operators have the same symmetries, and can be determined efficiently by solving a semidefinite programming problem.

An interesting future direction to pursue is to use the optimality conditions we developed in this paper to derive closed form solutions for other special cases.

APPENDIX A: PROOF OF Eq. (30)

To develop the optimal dual solution in the case of orthogonal null spaces, let $\Theta = [\Theta_1\Theta_2\cdots\Theta_m]$, and define a matrix $\Theta^\perp$ such that $[\Theta\ \Theta^\perp]$ is a square, unitary matrix, i.e., $[\Theta\ \Theta^\perp]^*[\Theta\ \Theta^\perp] = I$. Denoting $Z = [\Theta\ \Theta^\perp]Y[\Theta\ \Theta^\perp]^*$, the dual problem can be expressed as

$$\min_Y \mathrm{Tr}([\Theta\ \Theta^\perp]Y[\Theta\ \Theta^\perp]^*)\tag{A1}$$

subject to

$$\Theta_i^*[\Theta\ \Theta^{\perp}]Y[\Theta\ \Theta^{\perp}]^*\Theta_i \geqslant \Theta_i^* p_i\rho_i\Theta_{i,} \quad 1 \leqslant i \leqslant m;$$

$$Y \geqslant 0. \tag{A2}$$

Using the orthogonality properties of Θ_i and $\Theta^{\perp}$, the problem of Eqs. (A1) and (A2) is equivalent to

$$\min_Y \mathrm{Tr}(Y) \tag{A3}$$

subject to

$$Y_i \geqslant \Theta_i^* p_i\rho_i\Theta_i, \quad 1 \leqslant i \leqslant m;$$

$$Y \geqslant 0, \tag{A4}$$

where

$$Y = \begin{bmatrix} Y_1 & & & & \\ & Y_2 & & & \\ & & \ddots & & \\ & & & Y_m & \\ & & & & 0 \end{bmatrix}. \tag{A5}$$

Since $\mathrm{Tr}(Y)=\Sigma_{i=1}^m \mathrm{Tr}(Y_i)$, a solution to Eq. (A3) subject to Eq. (A4) is

$$\hat{Y} = \begin{bmatrix} \hat{Y}_1 & & & & \\ & \hat{Y}_2 & & & \\ & & \ddots & & \\ & & & \hat{Y}_m & \\ & & & & 0 \end{bmatrix}, \tag{A6}$$

where

$$\hat{Y} = \Theta_i^* p_i\rho_i\Theta_i, \quad 1 \leqslant i \leqslant m. \tag{A7}$$

Then,

$$\hat{Z} = [\Theta\ \Theta^{\perp}]\hat{Y}[\Theta\ \Theta^{\perp}]^* = \sum_{i=1}^m p_iP_i\rho_iP_i, \tag{A8}$$

as in Eq. (30).

APPENDIX B: PROOF OF Eq. (31)

To develop the optimal dual solution $\hat{Z}$ for one-dimensional null spaces, we note that $\hat{Z}$ lies in the space spanned by Θ_1 and Θ_2. Denoting by Θ a matrix whose columns represent an orthonormal basis for this space, $\hat{Z}$ can be written as $\hat{Z}=\Theta\hat{Y}\Theta^*$, where the 2×2 matrix $\hat{Y}$ is the solution to

$$\min_Y \mathrm{Tr}(Y) \tag{B1}$$

subject to

$$\Phi_1^* Y\Phi_1 \geqslant d_1, \tag{B2}$$

$$\Phi_2^* Y\Phi_2 \geqslant d_2, \tag{B3}$$

$$Y \geqslant 0. \tag{B4}$$

Here $\Phi_i=\Theta^*\Theta_i$ and $d_i=p_i\Theta_i^*\rho_i\Theta_i$ for $1\leqslant i\leqslant 2$.

To develop a solution to Eq. (B1) subject to Eqs. (B2)–(B4), we form the Lagrangian

$$\mathcal{L} = \mathrm{Tr}(Y) - \sum_{i=1}^2 \gamma_i(\Phi_i^*Y\Phi_i - d_i) - \mathrm{Tr}(XY), \tag{B5}$$

where from the Karush-Kuhn-Tucker conditions [32] we must have that $\gamma_i \geqslant 0, X \geqslant 0$, and

$$\gamma_i(\Phi_i^*Y\Phi_i - d_i) = 0, \quad i = 1,2, \tag{B6}$$

$$\mathrm{Tr}(XY) = 0. \tag{B7}$$

Differentiating $\mathcal{L}$ with respect to Y and equating to zero,

$$I - \sum_{i=1}^2 \gamma_i\Phi_i\Phi_i^* - X = 0. \tag{B8}$$

If $X=0$, then we must have that $I=\Sigma_{i=1}^2\gamma_i\Phi_i\Phi_i^*$, which is possible only if Φ_1 and Φ_2 are orthogonal. Therefore, $X \neq 0$, which implies from Eq. (B7) that Eq. (B4) is active. Now, suppose that only Eq. (B4) is active. In this case our problem reduces to minimizing $\mathrm{Tr}(y^*y)$ whose optimal solution is $y=0$, which does not satisfy Eqs. (B2) and (B3).

We conclude that at the optimal solution (B4) at least one of the constraints (B2) and (B3) is active. Thus, to determine the optimal solution we need to determine the solutions under each of the three possibilities: only Eq. (B2) is active, only Eq. (B3) is active, both Eqs. (B2) and (B3) are active, and then choose the solution with the smallest objective.

Consider first the case in which Eqs. (B2) and (B4) are active. In this case, $\hat{Y}=\hat{y}\hat{y}^*$ for some vector $\hat{y}$, and without loss of generality we can assume that

$$\Phi_1^*\hat{y} = d_1. \tag{B9}$$

To satisfy Eq. (B9), $\hat{y}$ must have the form

$$\hat{y} = \sqrt{d_1}\Phi_1 + \hat{s}\Phi_1^{\perp}, \tag{B10}$$

where $\Phi_1^{\perp}$ is a unit norm vector orthogonal to Φ_1, so that $\Phi_1^*\Phi_1^{\perp}=0$, and $\hat{s}$ is chosen to minimize $\mathrm{Tr}(\hat{Y})$. Since

$$\mathrm{Tr}(\hat{Y}) = \hat{y}^*\hat{y} = d_1 + |\hat{s}|^2, \tag{B11}$$

$\hat{s}=0$. Thus, $\hat{Y}=d_1\Phi_1\Phi_1^*$, and $\mathrm{Tr}(\hat{Y})=d_1$. This solution is valid only if Eq. (B3) is satisfied, i.e., only if

$$\Phi_2^*\hat{Y}\Phi_2 = d_1|f|^2 \geqslant d_2. \tag{B12}$$

Here we used the fact that

$$\Phi_2^*\Phi_1 = \Theta_2^*\Theta\Theta^*\Theta_1 = \Theta_2^*\Theta_1 = f, \tag{B13}$$

since $\Theta\Theta^*$ is an orthogonal projection onto the space spanned by Θ_1 and Θ_2.

Next, suppose that Eqs. (B3) and (B4) are active. In this case, $\hat{Y}=\hat{y}\hat{y}^*$ where without loss of generality we can choose $\hat{y}$ such that

$$\Phi_2^*\hat{y} = d_2 \tag{B14}$$

and

$$\hat{y} = \sqrt{d_2}\Phi_2 + \hat{s}\Phi_2^{\perp}, \tag{B15}$$

where $\Phi_2^{\perp}$ is a unit norm vector orthogonal to Φ_2, and $\hat{s}$ is chosen to minimize $\mathrm{Tr}(\hat{Y})$. Since $\mathrm{Tr}(\hat{Y})=d_2+|\hat{s}|^2$, $\hat{s}=0$, and $\mathrm{Tr}(\hat{Y})=d_2$. This solution is valid only if Eq. (B2) is satisfied, i.e.,

$$\Phi_1^* Y \Phi_1 = d_2|f|^2 \geqslant d_1. \tag{B16}$$

Finally, consider the case in which Eqs. (B2)–(B4) are active. In this case, we can assume without loss of generality that $\Phi_2^*\hat{y}=\sqrt{d_2}$. Then,

$$\hat{y} = \sqrt{d_2}\Phi_2 + \hat{s}\Phi_2^{\perp}, \tag{B17}$$

where $\hat{s}$ is chosen such that

$$\Phi_1^*\hat{Y}\Phi_1 = d_1 \tag{B18}$$

and $\mathrm{Tr}(\hat{Y})=\hat{y}^*\hat{y}$ is minimized. Now, for $\hat{y}$ given by Eq. (B17),

$$\hat{Y} = d_2\Phi_2\Phi_2^* + |\hat{s}|^2\Phi_2^{\perp}\Phi_2^{\perp *} + \hat{s}\sqrt{d_2}\Phi_2^{\perp}\Phi_2^* + \hat{s}^*\sqrt{d_2}\Phi_2\Phi_2^{\perp *}, \tag{B19}$$

so that

$$\begin{aligned}\Phi_1^*\hat{Y}\Phi_1 &= d_2|f|^2 + |\hat{s}|^2|e|^2 + \sqrt{d_2}\hat{s}e^*f + \sqrt{d_2}\hat{s}^*f^*e\\ &= |\sqrt{d_2}f + \hat{s}^*e|^2,\end{aligned} \tag{B20}$$

where we defined $\Theta_2^{\perp}=\Theta\Psi_2^{\perp}$, and e and f are given by Eq. (32). Therefore, to satisfy Eq. (B18), $\hat{s}$ must be of the form

$$\hat{s} = \frac{1}{e^*}(e^{j\varphi}\sqrt{d_1} - f^*\sqrt{d_2}) \tag{B21}$$

for some φ. The problem of Eq. (B1) then becomes

$$\min_{\varphi}\frac{1}{|e|^2}|e^{j\varphi}\sqrt{d_1} - f^*\sqrt{d_2}|^2, \tag{B22}$$

which is equivalent to

$$\max_{\varphi}\mathrm{Re}\{e^{j\varphi}f\}. \tag{B23}$$

Since

$$\mathrm{Re}\ \{e^{j\varphi}f\} \leqslant |e^{j\varphi}f| = |f|, \tag{B24}$$

the optimal choice of φ is $e^{j\varphi}=f^*/|f|$, and

$$\hat{s} = \frac{f^*\sqrt{d_2}}{e^*}\left(\frac{\sqrt{d_1}}{\sqrt{d_2}|f|} - 1\right). \tag{B25}$$

For this choice of $\hat{s}$,

$$\mathrm{Tr}(\hat{Y}) = d_2 + |\hat{s}|^2 = d_2\left[1 + \frac{|f|^2}{|e|^2}\left(\frac{\sqrt{d_1}}{\sqrt{d_2}|f|} - 1\right)^2\right] \triangleq \alpha. \tag{B26}$$

Clearly, $\alpha\geqslant d_2$. Therefore, to complete the proof of Eq. (31) we need to show that $\alpha\geqslant d_1$. Now,

$$\begin{aligned}|e|^2(\alpha - d_1) &= |e|^2(d_2 - d_1) + |f|^2\left(\frac{\sqrt{d_1}}{|f|} - \sqrt{d_2}\right)^2\\ &= (1 - |e|^2)d_1 + (|e|^2 + |f|^2)d_2 - 2\sqrt{d_1}\sqrt{d_2}|f|\\ &= (|f|\sqrt{d_1} - \sqrt{d_2})^2 \geqslant 0,\end{aligned} \tag{B27}$$

where we used the fact that

$$|e|^2 + |f|^2 = \Theta_1^*\Theta_2\Theta_2^*\Theta_1 + \Theta_1^*\Theta_2^{\perp}(\Theta_2^{\perp})^*\Theta_1 = \Theta_1^*\Theta_1 = 1, \tag{B28}$$

since $\Theta_2\Theta_2^*+\Theta_2^{\perp}(\Theta_2^{\perp})^*$ is an orthogonal projection onto the space spanned by Θ_1 and Θ_2.

[1] A. S. Holevo, J. Multivariate Anal. **3**, 337 (1973).

[2] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[3] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **49**, 1012 (2003).

[4] M. Charbit, C. Bendjaballah, and C. W. Helstrom, IEEE Trans. Inf. Theory **35**, 1131 (1989).

[5] M. Osaki, M. Ban, and O. Hirota, Phys. Rev. A **54**, 1691 (1996).

[6] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).

[7] Y. C. Eldar and C. D. Forney, Jr., IEEE Trans. Inf. Theory **47**, 858 (2001).

[8] Y. C. Eldar, A. Megretski, and G. C. Verghese, e-print quant-ph/0211111.

[9] Y. C. Eldar, Phys. Rev. A **68**, 052303 (2003).

[10] H. P. Yuen, R. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **IT-21**, 125 (1975).

[11] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987).

[12] D. Dieks, Phys. Lett. A **126**, 303 (1988).

[13] A. Peres, Phys. Lett. A **128**, 19 (1988).

[14] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).

[15] A. Peres and D. R. Terno, J. Phys. A **31**, 7105 (1998).

[16] A. Chefles, Phys. Lett. A **239**, 339 (1998).

[17] A. Chefles and S. M. Barnett, Phys. Lett. A **250**, 223 (1998).

[18] Y. C. Eldar, IEEE Trans. Inf. Theory **49**, 446 (2003).

[19] Y. C. Eldar, Phys. Rev. A **67**, 042309 (2003).

[20] C. W. Zhang, C. F. Li, and G. C. Quo, Phys. Lett. A **261**, 25 (1999).

[21] J. Fiurášek and M. Ježek, Phys. Rev. A **67**, 012321 (2003).

[22] T. Rudolph, R. W. Spekkens, and P. S. Turner, Phys. Rev. A **68**, 010301 (2003).

[23] P. Raynal, N. Lutkenhaus, and S. J. van Enk, Phys. Rev. A **68**, 022308 (2003).

[24] A. Peres, Found. Phys. **20**, 1441 (1990).

[25] Y. C. Eldar and H. Bölcskei, IEEE Trans. Inf. Theory **49**, 993 (2003).

[26] F. Alizadeh, Ph.D. thesis, University of Minnesota, Minneapolis, MN 1991 (unpublished).

[27] F. Alizadeh, in *Advances in Optimization and Parallel Computing*, edited by P. Pardalos (North-Holland, Netherlands, 1992).

[28] Y. Nesterov and A. Nemirovski, *Interior-Point Polynomial Algorithms in Convex Programming* (SIAM, Philadelphia, 1994).

[29] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 40 (1996).

[30] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. (Athena Scientific, Belmont, MA, 1999).

[31] G. D. Forney, Jr., IEEE Trans. Inf. Theory **37**, 1241 (1991).

[32] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization* (SIAM, Philadelphia, 2001).

[33] Otherwise we can transform the problem to a problem equivalent to the one considered in this paper by reformulating the problem on the subspace spanned by the eigenvectors of $\{\rho_i, 1 \leqslant i \leqslant m\}$.

[34] Interior point methods are iterative algorithms that terminate once a prespecified accuracy has been reached. A worst-case analysis of interior point methods shows that the effort required to solve a semidefinite program to a given accuracy grows no faster than a polynomial of the problem size. In practice, the algorithms behave much better than predicted by the worst-case analysis, and in fact in many cases the number of iterations is almost constant in the size of the problem.